International Journal of Technology Assessment in Health Care, **21:3** (2005), 298–304.

Self-reports of health-care utilization: Diary or questionnaire?

Mandy van den Brink, Wilbert B. van den Hout, Anne M. Stiggelbout, Hein Putter, Cornelis J. H. van de Velde, Job Kievit

Leiden University Medical Center

Objectives: The feasibility and convergent validity of a cost diary and a cost questionnaire was investigated.
Methods: Data were obtained as part of a cost-utility analysis alongside a multicenter clinical trial in patients with resectable rectal cancer. A sample of 107 patients from 30 hospitals was asked to keep a weekly diary during the first 3 months after surgery, and a monthly diary from 3 to 12 months after surgery. A second sample of seventy-two patients from twenty-eight hospitals in the trial received a questionnaire at 3, 6, and 12 months after surgery, referring to the previous 3 or 6 months. Format and items of the questions were similar and included a wide range of medical and nonmedical items and costs after hospitalization for surgery.
Results: Small differences were found with respect to nonresponse (range, 79 to 86 percent) and missing questions (range, 1 to 6 percent between the diary and questionnaire). For most estimates of volumes of care and of costs, the diary and questionnaire did not differ significantly. Total 3-month nonhospital costs were €1,860, €1,280, and €1,050 in the diary sample and €1,860, €1,090, and €840 in the questionnaire sample at 3, 6, and 12 months after surgery, respectively ($p = .50$). However, with respect to open questions, the diary sample tended to report significantly more care.
Conclusions: For the assessment of health-care utilization in economic evaluations alongside clinical trials, a cost questionnaire with structured closed questions may replace a cost diary for recall periods up to 6 months.

Keywords: Economic evaluation, Clinical trial, Cost measurement, Self-report, Methodology, Rectal cancer

Economic evaluations in health care, performed from a societal perspective, include the measurement of medical and nonmedical costs (7). For many cost items, costs can be estimated as volumes times prices. For prices, one may revert to cost calculations, standard prices, or charges. For volumes, estimates from other studies are most frequently used (19). However, these estimates cannot always be found for all types of care and may not apply to other settings (1;12). To obtain more-valid volume estimates, reports from providers or recipients of care often need to be used.

Providers of care can be expected to give the most accurate and detailed information on the type and volumes of care by using their administrative systems. However, the potential multiplicity of care providers may limit the feasibility of obtaining data at the patient level. Alternatively, by means of self-report, care from multiple providers can be assessed directly from the patient. Methods of self-report are questionnaires, diaries, and interviews. Compared with diaries and interviews, a retrospective questionnaire is less labor-intensive, thus less costly, and requires less motivation

This study was supported by grant 97-026 from the Health Care Insurance Board, The Hague, The Netherlands, and by grant 94-10-036 from the Netherlands Organisation for Health Research and Development, The Hague, The Netherlands. The authors also gratefully acknowledge the Dutch Colorectal Cancer Group and other cooperative investigators of the TME study for their participation in this study. The results of this study, in part, have been presented at the 16th annual meeting of the International Society for Technology Assessment in Health Care, 2000, The Hague, The Netherlands.

of patients and researchers. In the more-burdensome diary, selective nonresponse and missing values might occur more often, for example with increasing data collection periods, increasing patient age, or deteriorating patient health (2;17;18). On the other hand, a diary method may reduce recall error and, therefore, lead to data that are more valid (8;14). Few studies have compared methods of self-report for health-care utilization in patients. Goossens et al. (9) studied the use of a weekly cost diary in chronic back pain patients for different data collection periods. Extrapolated yearly costs did not differ significantly between the data collection periods, because the costs were stable over time for this type of chronic disease. The impact of different recall periods on the validity of the results was not studied, whereas this is one of the main considerations when choosing between a questionnaire and a diary method.

In other contexts, for example in public health surveys and quality of life research, self-report methods have been compared more extensively. Self-report, compared with medical records, then appears to be more accurate for more-important and less-frequent events (e.g., hospitalizations), for younger respondents, and for people in better health (5;6;10;20). Indications were found that data quality was higher for interviews than for mailed questionnaires (4;16), although in interviews, patients may tend to give more socially acceptable rather than true answers (23). Most studies concluded, however, that there were no or only small differences between methods of self-report (4;10;11;21;22;24). However, whether these results also apply to the assessment of health-care utilization in specific patient groups needs to be established. The purpose of this study was to assess the feasibility and convergent validity of questionnaire and diary methods for the measurement of health-care utilization in patients participating in a clinical trial, for different recall periods.

METHODS

Design and Subjects

Data were obtained as part of a cost-utility analysis alongside a randomized clinical trial (3;13). The main objective of this trial was to assess the additional value of preoperative radiotherapy (PRT) to total mesorectal excision (TME) in patients diagnosed with rectal cancer. Between January 1996 and January 2000, a total of 1,530 Dutch patients with resectable rectal cancer from 84 hospitals were randomized for TME surgery with or without PRT.

From February 1999 to January 2000, forty hospitals (48 percent) gave (medical ethical) approval to approach their patients for participation in utility interviews and a cost diary (*diary sample*). Reasons for withholding approval were, for example, the expected burden of the utility interviews for patients, no time to start up the medical ethical procedure, an expected small number of patients, too many other studies, or the nearby closing of the TME study. Patients in the diary sample were asked to participate in utility interviews just before treatment and at 3 (T3) and 12 months (T12) after surgery, and to fill out cost forms weekly from discharge to T3 and monthly from T3 to T12. Participants received a binder containing the cost forms during the interview before treatment and were encouraged to fill out the questions prospectively (i.e., to record volumes of care immediately after use), but at least weekly or monthly. Compliance phone calls were made at 1.5 and 8 months after TME surgery to encourage further completion and minimize unclear answers and missing values.

Asking one sample of patients to fill out both a diary and a retrospective questionnaire would allow for within-subject comparison but could introduce attention bias in the questionnaire due to the preceding diary. Therefore, a second sample of patients was asked to fill out a mailed cost form, along with the regular quality of life questionnaires at 3, 6 (T6), and 12 months after surgery, referring to the previous 3 (T3, T6) and 6 (T12) months (*questionnaire sample*). From April 1999 to January 2000, all TME patients who were not approached for participation in the diary sample were included in the questionnaire sample. If a cost form was not returned twice, the subsequent forms were not sent. No reminders were sent in case of nonresponse.

Instruments

The cost form included structured no/yes questions on medical and nonmedical care, each consisting of multiple items. Questions were asked on hospitalizations, contacts with general practitioners; outpatient visits to paramedics, radiotherapists, surgeons, internists, and other health-care workers; home help, district nursing, and informal care; and on medications, special food, care products, assistive devices, and miscellaneous items. If patients indicated that they had received specific care, they were asked for the volume: number of hospitalized days, number of contacts, number of outpatient visits, hours of care, and the types of obtained medications, special food, care products, assistive devices, and miscellaneous items, respectively. The cost forms were preceded by an instruction with examples for each question and a phone number to dial in case of questions. Patients were instructed to record all health-related care, for nonhospitalized periods only. The clarity and completeness of the cost form was pretested in a pilot study and adapted accordingly. Data on age, gender, adjuvant treatment, and perceived health, measured by a 100-mm visual analogue scale, ranging from 0 (death) to 100 (perfect health), were derived from the general TME study database.

Analyses

Coding of Data. In the diary sample, respondents were included only if at least two thirds of the cost forms of each data collection period (discharge to T3, T3–T6, and

https://doi.org/10.1017/S0266462305050397 Published online by Cambridge University Press

T6–T12) was filled out. Missing questions were replaced by the mean of the subject's nonmissing answers. In the questionnaire sample, respondents were included only if the cost form was returned with at least one filled out question. Missing questions were replaced by the mean of the other respondents in the questionnaire sample.

The number of hours of informal care was truncated to at most 4 hours per day. The number of medications, special food products, care products, assistive devices, and miscellaneous items were calculated by counting the number of different products, irrespective of the volume. In total, 927 distinct descriptions were identified and recoded to 233 types of medications, 42 types of special food, 148 types of care products, 21 types of assistive devices, and 32 types of miscellaneous items. Ten percent of the answers could not be identified and were excluded from the analyses.

If available, costs of medication, special food, care products, assistive devices, and miscellaneous items were calculated using standard daily doses, as recommended by the Dutch National Health Authority. Otherwise, information on standard use and cost prices were retrieved from the internet (www.medicijnen.net), and from suppliers of stoma care products. Reported volumes of hospital days, contacts with general practitioners, outpatient visits, and hours of care were multiplied by standard cost prices (15). All costs were estimated for the year 2002 and are presented in Euros.

Feasibility. To assess feasibility, response rates and the number of missing values were compared between the diary and questionnaire method using Chi-squared tests and Mann–Whitney tests. Logistic regression analyses were performed to investigate whether patient characteristics were related to nonresponse. The respondents included in the analyses in both samples were compared by age, gender, PRT, hospital stay, and health, using *t*-tests and Chi-squared statistics.

Convergent Validity. To assess convergent validity, we compared the number of hospital days, contacts with general practitioners, outpatient visits, hours of care, medication types, special food products, care products, assistive devices, miscellaneous items, and total nonhospital costs as reported in the diary and questionnaire during each data collection period (T0–T3, T3–T6, T6–T12). Reported volumes of hospital days, contacts with general practitioners, outpatient visits, hours of care, and nonhospital costs were standardized to 3-month volumes and costs for comparison over time.

Differences between the diary and questionnaire method were evaluated using repeated measures analyses with time as within-subjects and method as between-subjects factor, controlled for significant differences in patient characteristics between the samples. In case of a significant ($p < .05$) effect of method, post hoc analyses were carried out by data collection period. To study the sensitivity of our results to the assumptions of parametric methods, all analyses of volumes and costs were also carried out using log transformations.

Table 1. Patient Characteristics

	Diary (n = 107)	Questionnaire (n = 72)	*p* value
Characteristic			
Males (%)	63	75	.06
Mean age in years (SD)	63 (10)	63 (10)	.98
Perceived health: mean VAS score (SD)	73 (17)	71 (20)	.36
Preoperative radiotherapy (%)	51	68	.02
Mean no. of hospital days for TME surgery (SD)	21 (18)	19 (17)	.40

VAS, visual analog scale; TME, total mesorectal excision.

RESULTS

Subjects

In total, 169 patients were asked to participate in the diary sample, and 112 (66 percent) patients from thirty hospitals consented. Most refusals were due to the patient's belief that the accompanying utility interviews would be too burdensome (67 percent). One patient explicitly mentioned the diary as the reason not to participate. No statistically significant differences were found between participants and nonparticipants in the diary sample concerning PRT ($p = .25$), gender ($p = .87$), adjuvant treatment ($p = .38$), and perceived health ($p = .42$). However, participants were significantly younger (mean age, 63 years; SD 10) than nonparticipants (mean age, 67 years; SD 11; $p = .03$). Four patients died in hospital shortly after surgery. One patient was not operated on and was excluded from the analyses. Thus, the diary sample initially consisted of 107 patients.

Seventy-two patients from twenty-eight hospitals in the trial were included in the questionnaire sample, and mailed a cost form. The characteristics of the patients included in both samples are shown in Table 1.

In the questionnaire sample, significantly more patients had received PRT and were male. The difference in PRT was caused by a temporary change in the randomization procedure of the TME study at the time of the questionnaire study. The difference in gender may be explained by a difference in the baseline gender distribution. To account for these differences, we controlled for PRT and gender in all analyses.

Feasibility

Response rates did not differ significantly between the diary and questionnaire sample (Table 2). At T3, T6, and T12, response rates were 79 percent, 84 percent, and 83 percent for the diary and 86 percent, 85 percent, and 82 percent for the questionnaire, respectively ($p = .17$, $p = .65$, and $p = .54$, respectively). The somewhat lower response for the diary at T3 is mainly attributable to the fact that 20 patients (19 percent) in the diary sample were interviewed after hospital discharge, because, for reasons of logistics, they could not be interviewed before surgery. This resulted in a high

Table 2. Responses by Time and Method

	Discharge to 3 months		3 to 6 months		6 to 12 months	
	Diary N (%)	Questionnaire N (%)	Diary N (%)	Questionnaire N (%)	Diary N (%)	Questionnaire N (%)
Initial sample size[a]	107	72	107	71	105	68
No form(s) returned	13 (12)	10 (14)	12 (11)	11 (15)	13 (12)	12 (18)
More than 1/3 missing forms	9 (9)	–	5 (5)	–	5 (5)	–
Respondents analyzed	85 (79)	62 (86)	90 (84)	60 (85)	87 (83)	56 (82)

[a] The initial sample sizes decrease as a result of people who die during the course of the investigation.

number of missing forms in the first weeks after discharge (Figure 1). The percentage of missing forms in the patients interviewed before surgery remained relatively stable over time. PRT, age, perceived health, gender, adjuvant treatment, and sample were not significantly related to nonresponse at T3, T6, and T12.

The mean percentage of forms filled out by the respondents analyzed in the diary sample was 96 percent at T3 and 99 percent at T6 and T12, respectively. The mean number of missing questions was low at all times (6 percent, 2 percent, and 2 percent for the diary versus 2 percent, 1 percent, and 5 percent for the questionnaire at T3, T6, and T12, respectively), but significantly more missing questions were seen in the diary at T3 ($p < .001$) and T6 ($p = .01$). Analysis by type of question showed that the questions on contacts with general practitioners and outpatient visits to paramedics, medical specialists, and other health-care workers were missing more often in the diary sample ($p = .01$ and $p < .001$ at T3 and T6, respectively).

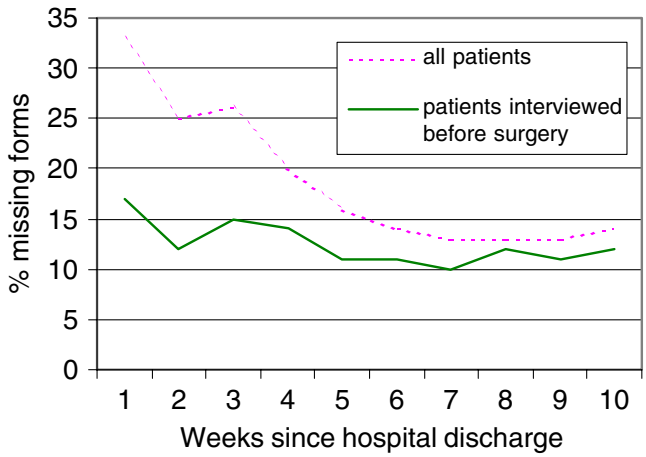

Figure 1. Missing weekly forms in the diary sample. The dotted line reflects the missing forms of all patients in the diary sample, either interviewed before or after hospitalization. The solid line depicts the missing forms of the patients that were interviewed before surgery. The percentage of missing forms is shown by time since hospital discharge, as people were asked to fill out the cost forms after hospitalization.

Convergent Validity

The estimated mean nonhospital costs did not differ significantly between the questionnaire and the diary method (Table 3, $p = .50$). Analysis by type of care showed no significant differences between the diary and questionnaire method for contacts with general practitioners; outpatient visits to paramedics and medical specialists; hours of home help, district nursing, and informal care; and types of special food and assistive devices. The diary sample did report more contacts with general practitioners and more hours of home help, but these differences disappeared when we controlled for differences in PRT and gender between the samples. A significant effect of measurement method was found for the number of hospital days ($p = .01$), the number of contacts with other health-care workers ($p < .001$), and the number of types of obtained medications, care products, and miscellaneous items (all $p < .05$).

Post hoc analyses by data collection period showed that significantly more hospital days (mean, 28; SD 73) were reported in the questionnaire than in the diary (mean, 6; SD 10; $p = .01$) only from T0 to T3. Significantly more contacts with other health-care workers were reported in the diary at all times (all $p < .001$). The number of types of care products and miscellaneous items was significantly larger in the diary from T0 to T3 (all $p < .01$). In addition, in the diary, more types of miscellaneous items and more types of medications were reported from T3 to T6 ($p = .001$) and T6 to T12 ($p = .01$), respectively. Re-analysis using log-transformed data did not substantially change any of the results (data not shown).

DISCUSSION

The aim of this study was to assess the feasibility and convergent validity of a questionnaire and a diary method for the measurement of health-care utilization in patients participating in a clinical trial. For this purpose, we compared the use of a cost diary and a cost questionnaire in two samples of patients. Comparisons were made in terms of response rates, sample representativeness, missing values, estimated costs and volumes of care, and by different recall periods.

https://doi.org/10.1017/S0266462305050397 Published online by Cambridge University Press

Table 3. Average Costs and Volumes (SD) of Care by Time and Method

	Discharge to 3 months		3 to 6 months		6 to 12 months		
	Diary (n = 85)	Questionnaire (n = 62)	Diary (n = 90)	Questionnaire (n = 60)	Diary (n = 87)	Questionnaire (n = 56)	*p* value[a]
Type of care							
No. of hospital days	6.0 (10)	28 (73)	2.3 (4.4)	2.2 (7.0)	1.4 (4.1)	1.2 (3.7)	.01
Contacts with							
General practitioners	4.9 (4.3)	4.1 (3.8)	2.0 (2.6)	1.7 (2.5)	1.5 (2.1)	1.1 (2.1)	.26
Paramedics	4.0 (5.1)	6.1 (9.0)	2.8 (8.6)	4.7 (19)	1.7 (4.4)	1.3 (3.7)	.32
Medical specialists	5.3 (5.0)	6.1 (7.2)	2.8 (2.8)	3.1 (3.0)	2.1 (2.2)	1.8 (1.5)	.67
Other health-care workers	2.4 (4.1)	0.2 (0.9)	1.2 (2.4)	0.1 (0.4)	0.9 (1.4)	0.2 (0.6)	.00
Hours of							
Home help	6.8 (20)	4.6 (13)	8.3 (19)	4.3 (17)	6.1 (16)	3.0 (11)	.44
District nursing	4.9 (10)	7.2 (20)	3.8 (13)	6.4 (37)	2.3 (10)	1.8 (8)	.55
Informal care	26 (60)	32 (75)	6.0 (17)	1.6 (8)	4.8 (14)	11 (51)	.80
Types of							
Medications	2.0 (2.0)	1.5 (1.5)	1.5 (1.8)	1.4 (1.5)	2.0 (2.1)	1.2 (1.4)	.04
Special food	0.4 (0.9)	0.3 (0.6)	0.1 (0.5)	0.2 (0.6)	0.1 (0.5)	0.2 (0.7)	.85
Care products	1.9 (1.7)	1.1 (0.8)	1.2 (1.0)	0.8 (0.7)	1.1 (1.3)	0.9 (0.8)	.01
Assistive devices	0.2 (0.6)	0.3 (0.7)	0.1 (0.5)	0.1 (0.5)	0.1 (0.4)	0.1 (0.7)	.72
Miscellaneous	0.4 (0.8)	0.0 (0.2)	0.3 (0.6)	0.0 (0.1)	0.2 (0.8)	0.1 (0.7)	.00
Nonhospital costs of care (€)	1,860 (1,100)	1,860 (1,350)	1,280 (1,290)	1,090 (1,560)	1,050 (1,290)	840 (980)	.50

[a] The analyses used to test the differences between measurement methods were controlled for differences in preoperative radiotherapy and gender between the samples. Note that the average volumes and costs shown in this table are uncorrected estimates.

For both measurement methods, response rates were high. At all times, the percentage of patients who did not return the cost forms was somewhat higher in the questionnaire sample. However, in the diary sample, additional patients were excluded because of incomplete diaries, resulting in response rates that did not differ significantly between the methods. Part of the nonresponse in the weekly diary was attributable to the fact that, for reasons of logistics, it had not been possible to instruct all patients in the diary sample before TME surgery. The administration of the cost diary without oral instruction would have solved this problem but might have resulted in more missing forms and more missing values, because patients would probably have been less motivated to complete the diary.

Total nonhospital costs did not differ significantly between the questionnaire and the diary method. Analysis by type of care also showed no differences for most volumes of care. During the interviews in the diary sample, many patients consulted a family member, or looked up their appointments in agendas when filling out the cost form. This mnemonic may also have helped respondents in the questionnaire sample to fill out the questions accurately and may explain in part why we found few differences between the methods. Significantly more hospital days were reported in the questionnaire from T0 to T3, probably because respondents erroneously included the days hospitalized for TME surgery. In later collection periods, the results on hospital days were very similar. In the diary, more contacts with other health-care workers, types of obtained medications, care products, and other expenses were reported. The length of the recall period was not strongly related to differences between the methods. Slightly more differences were found from T0 to T3 (i.e., weekly versus 3-month recall) than for the other data collection periods (i.e., monthly versus 3-month recall, and monthly versus 6-month recall). However, we did not compare all different recall periods concurrently. Therefore, we cannot exclude alternative hypotheses, for example, the frequency of events, and the motivation of respondents.

Remarkably, the observed differences in reported volumes of care between the methods all occurred in answers to open-ended questions. The repeated confrontation with open questions and the instruction of the interviewer may have motivated people in the diary sample to answer these questions more often. Only two open questions, that is, types of special food and assistive devices, did not result in significant differences between the diary and questionnaire method. For these types of care, the frequency of events may have been too low to reach significance. In addition to the problem of underreporting, open questions also provide problems in the interpretation and recoding of the answers. In paper questionnaires, it may not always be feasible to enumerate all possible answers. However, advances in technology, for example, the automated administration of questionnaires by means of a computer, may in the future solve this problem. The observed differences between the methods might also have been caused by differences in actual care the samples received, because we could not randomize patients. We do not consider this very likely, however. The patients in the questionnaire sample were retrieved mostly from hospitals that did not give approval to approach their patients for participation in

utility interviews and the cost diary. The reasons mentioned for not giving approval do not seem related to differences in provided care. In addition, the observed differences in gender and PRT between the samples were controlled for in all analyses.

In conclusion, we found only small differences between the diary and questionnaire method with respect to response rates, missing values, and for most estimates of nonhospital costs and volumes of care. Only for open questions, the diary sample tended to report significantly more care. Therefore, we conclude that, in economic evaluations alongside clinical trials, a cost questionnaire with structured closed questions may replace a cost diary for recall periods up to 6 months.

POLICY IMPLICATIONS

In economic evaluations alongside clinical trials, a cost questionnaire with structured closed questions is an efficient method to assess medical and nonmedical costs, as compared with a cost diary. Results from economic evaluations that have used cost questionnaires are not less valid than results from studies that have used cost diaries. Only for open questions, a cost diary may be a more-efficient method than a cost questionnaire, depending on the impact of the costs assessed by open questions on the results of the economic evaluation.

CONTACT INFORMATION

Mandy van den Brink, MSc, Research (m.vandenbrink@vumc.nl), **Wilbert B. van den Hout**, PhD, Assistant Professor (w.b.van_den_Hout@lumc.nl), **Anne M. Stiggelbout**, PhD, Associate Professor (a.m.stiggelbout@lumn.nl), Department of Medical Decision Making, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands

Hein Putter, PhD, Assistant Professor (h.putter@lumnc), Department of Medical Statistics, Leiden University Medical Center, Wassenaarseweg 62, 2333 AL Leiden, The Netherlands

Cornelis J. H. van de Velde, MD, PhD, Professor (c.j.h.van_de_velde@lumc.nl), Department of Surgery, **Job Kievit**, MD, PhD, Professor, (j.kievit@lumc.nl), Department of Medical Decision Making; Surgeon, Department of Surgical Oncology, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands

REFERENCES

1. Barber JA, Thompson SG. Analysis and interpretation of cost data in randomized controlled trials: Review of published studies. *BMJ*. 1998;317:1195-1200.
2. Biemer PP, Groves RM, Lyberg LE, et al. *Measurement errors in surveys*. New York: John Wiley & Sons; 1991.
3. van den Brink M, van den Hout WB, Stiggelbout AM, et al. Cost-utility analysis of pre-operative radiotherapy in patients with rectal cancer undergoing total mesorectal excision. *J Clin Oncol*. 2004;22:244-253.
4. van Campen C, Sixma H, Kerssens JJ, et al. Comparisons of the costs and quality of patient data collection by mail versus telephone versus in-person interviews. *Eur J Public Health*. 1998;8:66-70.
5. Carsjo K, Thorslund M, Warneryd B. The validity of survey data on utilization of health and social services among the very old. *J Gerontol*. 1994;49:S156-S164.
6. Cleary PD, Jette AM. The validity of self-reported physician utilization measures. *Med Care*. 1984;22:796-803.
7. Gold MR, Siegel JE, Russell LB, et al. *Cost-effectiveness in health and medicine*. New York: Oxford University Press; 1996.
8. Coughlin SS. Recall bias in epidemiologic studies. *J Clin Epidemiol*. 1990;43:87-91.
9. Goossens ME, Rutten-van Molken MP, Vlaeyen JW, et al. The cost diary: A method to measure direct and indirect costs in cost- effectiveness research. *J Clin Epidemiol*. 2000;53:688-695.
10. Green S, Kaufert J, Corkhill R, et al. The collection of service utilisation data: A research note on validity. *Soc Sci Med*. 1979;13A:231-234.
11. Hebert R, Bravo G, Korner-Bitensky N, et al. Refusal and information bias associated with postal questionnaires and face-to-face interviews in very elderly subjects. *J Clin Epidemiol*. 1996;49:373-381.
12. van den Hout WB, van den Brink M, Stiggelbout AM, et al. Cost-effectiveness analysis of colorectal cancer treatments. *Eur J Cancer*. 2002;38:953-963.
13. Kapiteijn E, Marijnen CA, Nagtegaal ID, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. *N Engl J Med*. 2001;345:638-646.
14. Lippman A, Mackenzie SG. What is "recall bias" and does it exist? *Prog Clin Biol Res*. 1985;163C:205-209.
15. Oostenbrink JB, Koopmanschap MA, Rutten FFH. Manual for cost analyzes, methods and standard prices for economic evaluations in health care (In Dutch). Amstelveen: Dutch Health Insurance Executive Board; 2000.
16. Perkins JJ, Sanson-Fisher RW. An examination of self- and telephone-administered modes of administration for the Australian SF-36. *J Clin Epidemiol*. 1998;51:969-973.
17. Ross MM, Rideout EM, Carson MM. The use of the diary as a data collection technique. *West J Nurs Res*. 1994;16:414-425.
18. Stone AA, Shiffman S, Schwartz JE. Patient non-compliance with paper diaries. *BMJ*. 2002;324:1193-1194.
19. Stone PW, Chapman RH, Sandberg EA, et al. Measuring costs in cost-utility analyzes. Variations in the literature. *Int J Technol Assess Health Care*. 2000;16:111-124.
20. Wallihan DB, Stump TE, Callahan CM. Accuracy of self-reported health services use and patterns of care among urban older adults. *Med Care*. 1999;37:662-670.
21. Weinberger M, Nagle B, Hanlon JT, et al. Assessing health-related quality of life in elderly outpatients: Telephone versus

face-to-face administration. *J Am Geriatr Soc*. 1994;42:1295-1299.
22. van Wijck EE, Bosch JL, Hunink MGM. Time-tradeoff values and standard-gamble utilities assessed during telephone interviews versus face-to-face interviews. *Med Decis Making*. 1998;18:400-405.
23. Wilson K, Roe B, Wright L. Telephone or face-to-face interviews?: A decision made on the basis of a pilot study. *Int J Nurs Stud*. 1998;35:314-321.
24. Wu AW, Jacobson DL, Berzon RA, et al. The effect of mode of administration on medical outcomes study health ratings and EuroQol scores in AIDS. *Qual Life Res*. 1997;6:3-10.